




A New Pneumococcal Capsule Type, 10D, is the 100th Serotype and Has a Large *cps* Fragment from an Oral Streptococcus

Feroze Ganaie,^a Jamil S. Saad,^b Lesley McGee,^c Andries J. van Tonder,^d Stephen D. Bentley,^e  Stephanie W. Lo,^e Rebecca A. Gladstone,^e Paul Turner,^{f,g} Jeremy D. Keenan,^h Robert F. Breiman,ⁱ Moon H. Nahm^a

^aDepartment of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA

^bDepartment of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, USA

^cRespiratory Diseases Branch, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

^dDepartment of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

^eParasites and Microbes, Wellcome Sanger Institute, Hinxton, Cambridge, United Kingdom

^fCambodia Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia

^gCentre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

^hDepartment of Ophthalmology, University of California, San Francisco, California, USA

ⁱEmory Global Health Institute, Emory University, Atlanta, Georgia, USA

ABSTRACT *Streptococcus pneumoniae* (pneumococcus) is a major human pathogen producing structurally diverse capsular polysaccharides. Widespread use of highly successful pneumococcal conjugate vaccines (PCVs) targeting pneumococcal capsules has greatly reduced infections by the vaccine types but increased infections by nonvaccine serotypes. Herein, we report a new and the 100th capsule type, named serotype 10D, by determining its unique chemical structure and biosynthetic roles of all capsule synthesis locus (*cps*) genes. The name 10D reflects its serologic cross-reaction with serotype 10A and appearance of cross-opsonic antibodies in response to immunization with 10A polysaccharide in a 23-valent pneumococcal vaccine. Genetic analysis showed that 10D *cps* has three large regions syntenic to and highly homologous with *cps* loci from serotype 6C, serotype 39, and an oral streptococcus strain (*S. mitis* SK145). The 10D *cps* region syntenic to SK145 is about 6 kb and has a short gene fragment of *wciN α* at the 5' end. The presence of this nonfunctional *wciN α* fragment provides compelling evidence for a recent interspecies genetic transfer from oral streptococcus to pneumococcus. Since oral streptococci have a large repertoire of *cps* loci, widespread PCV usage could facilitate the appearance of novel serotypes through interspecies recombination.

IMPORTANCE The polysaccharide capsule is essential for the pathogenicity of pneumococcus, which is responsible for millions of deaths worldwide each year. Currently available pneumococcal vaccines are designed to elicit antibodies to the capsule polysaccharides of the pneumococcal isolates commonly causing diseases, and the antibodies provide protection only against the pneumococcus expressing the vaccine-targeted capsules. Since pneumococci can produce different capsule polysaccharides and therefore reduce vaccine effectiveness, it is important to track the appearance of novel pneumococcal capsule types and how these new capsules are created. Herein, we describe a new and the 100th pneumococcal capsule type with unique chemical and serological properties. The capsule type was named 10D for its serologic similarity to 10A. Genetic studies provide strong evidence that pneumococcus created 10D capsule polysaccharide by capturing a large genetic fragment from an oral streptococcus. Such interspecies genetic exchanges could greatly increase diversity of pneumococcal capsules and complicate serotype shifts.

KEYWORDS capsule, genetic exchange, *Streptococcus pneumoniae*, vaccine

Citation Ganaie F, Saad JS, McGee L, van Tonder AJ, Bentley SD, Lo SW, Gladstone RA, Turner P, Keenan JD, Breiman RF, Nahm MH. 2020. A new pneumococcal capsule type, 10D, is the 100th serotype and has a large *cps* fragment from an oral streptococcus. mBio 11:e00937-20. <https://doi.org/10.1128/mBio.00937-20>.

Editor Larry S. McDaniel, University of Mississippi Medical Center

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Moon H. Nahm, mnahm@uabmc.edu.

This article is a direct contribution from Moon H. Nahm, a Fellow of the American Academy of Microbiology, who arranged for and secured reviews by Jason Rosch, St. Jude Children's Research Hospital, and Herve Tettelin, University of Maryland School of Medicine.

Received 16 April 2020

Accepted 20 April 2020

Published 19 May 2020

Streptococcus pneumoniae (pneumococcus) has a storied history. It is one of the first pathogens ever discovered and is a major cause of pneumonia and mortality (1, 2). Its virulence is largely due to its thick polysaccharide (PS) capsule (3), and studies of its capsule elucidated that the genetic material was DNA (4). Currently, the capsule is used to produce pneumococcal conjugate vaccines (PCVs), a major public health tool worldwide. Since PCV protection is serotype specific, PCVs provide protection only against the serotypes included in the vaccine (5). Pneumococcus displays one of many structurally diverse capsules, with 99 identified so far (Table 1) (6). The widespread use of PCVs has altered the natural distribution of pneumococcal serotypes by increasing invasive pneumococcal diseases (IPDs) and nasopharyngeal carriage of nonvaccine serotypes (7). This change in serotype distribution is called serotype replacement (8–10). Diversity in pneumococcal capsule types and serotype replacement are major challenges in making and using PCVs.

Serotype replacement can occur in several ways. It may occur when a minor preexisting serotype increases its prevalence following vaccination. For instance, PCV vaccination could have increased the prevalence of serotypes 6C and 19A, which previously existed as minor populations (11–14). Alternatively, replacement may occur if a pneumococcus strain expressing a new capsule type appears by capturing a gene fragment from another pneumococcal strain or related species like oral streptococcus (15). The genetic transfer is facilitated since almost all the genes necessary for capsule biosynthesis are in a genetic cassette called *cps* in these species. Knowledge of capsular PS diversity, and understanding the origin of capsule diversity, is of fundamental importance in using PCVs and controlling pneumococcal infections (16).

To investigate capsule diversity, capsule loci of 21,853 pneumococcal genomes were studied as a part of the Global Pneumococcal Sequencing (GPS) project (17) and a mother-infant carriage study in the Maela refugee camp, Thailand (18). The genome studies revealed that several nasopharyngeal isolates from the children of Ethiopia (19), Cambodia (20), and Thailand (21) express a novel type of *cps*, labeled 39X. Their *cps* loci have parts of serotypes 6C and 39, separated by a gene of unknown origin, *wcrO*. The isolates were previously serotyped as either 39 or 10A by the Quellung and latex agglutination tests (17, 19–21). These isolates were provisionally designated as expressing serotype 39X, pending biochemical characterization, since a large genetic difference in *cps* may not confer a difference in the PS (22). Here, we determined the chemical structure of the 39X capsule to be unique, named the new type serotype 10D, and show that its *cps* has captured a large genetic fragment from oral streptococci.

RESULTS

Serotype 10D (39X) displays unique serologic properties. Following the confirmation that a pneumococcal strain, Cam853, has the genetic marker described for a novel capsule type “39X” (21) (see Fig. S1 in the supplemental material), we investigated its serologic properties along with five strains expressing related serotypes (serotypes 39, 10A, 10B, 10C, and 10F) (23) using a panel of commercially available antisera and agglutination reaction (Table 2). Antisera used include factor serum, type serum, group serum, and pool serum, which recognize an epitope, serotype, serogroup, and multiple serotypes, respectively. Strain Cam853 agglutinated with factor serum 10d (FS10d) but failed to react with other antisera, including type 39 antiserum, even though the antiserum reacted with the related strains as expected (23). In addition, Cam853 reproducibly failed to agglutinate with any of the three pool antisera (pool E, S, and T) as well as with serogroup 6 reagents. Therefore, Cam853 appeared to have serologic properties not described so far for pneumococci.

To investigate its serologic properties in detail, we examined strain Cam853 along with the five strains expressing related serotypes by using flow cytometry, which is more sensitive and reproducible than agglutination reactions. Reactions with type 39 antiserum and a monoclonal antibody (MAB) Hyp10AG1, which was produced as described previously (23), distinguished strain Cam853 and serotypes 39 and 10A from other related serotypes (10B, 10C, 10F, and 6C) (Fig. 1 and Table 3). Strain Cam853 could

TABLE 1 List of the last 10 pneumococcal serotypes discovered with new serotyping approaches

Serotype no.	Serotype name	Chemical structure	Accession no. for <i>cps</i>	Year	Reference
91	6C	6C has repeating units of 6A, except the galactose residue is replaced with a glucose residue	EF538714	2007	11
92	11E	11E has repeating units of 11A except O-acetylation of a 1-phosphoglycerol	GU074953	2010	54
93	20B	20B has repeating unit of 20A with an extra branching glucose residue	JQ653093	2012	55
94	6D	6D has repeating units of 6C except rhamnose-(1→4)-ribitol linkage	HM171374	2013	46
95	6F	6F has both 6A and 6C repeating units	KC832410	2013	46
96	6G	6G has both 6B and 6D repeating units	KC832411	2013	46
97	6H	6H has both 6A and 6B repeating units	KJ874439	2015	56
98	35D	35D has repeating units of 35B without O-acetyl group at one Galf	KY084476	2017	46
99	7D	7D has 5:1 combination of 7C and 7B repeating units	NA ^a	2018	57
100	10D	See Fig. 4	ERR051587	2019	This study

^aNA, not available.

be further distinguished from serotypes 10A and 39 by their reactions with pool S and group 10 antisera. Cam853 reacted with pool S, but not with group 10 antiserum, whereas serotype 39 reacted with none and serotype 10A reacted with both (Fig. 1). Additional serologic properties are summarized in Table 3. The other two 39X strains (PATH4346 and Cam657) showed the same serologic profiles as Cam853 (Fig. S2). Taken together, 39X shares epitopes associated with serotype 10A but has a unique serologic profile.

The provisional name 39X was replaced with a new name, 10D, for the following reasons. Although 39X is genetically most similar to 39, it has substantial similarity to 6C and 10A *cps*. Serologically, 39X cross-reacts with FS10d and Hyp10AG1, which recognize a structural motif shared by serotypes 39 and 10A (Fig. S3) (24). Indeed, Cam853 was mistyped as 10A (17, 20). Further, 10A will be in a future PCV (25) and may elicit antibodies cross-reacting with 39X. Thus, the name 10D would facilitate cross-protection investigations of new PCVs.

10D capsular polysaccharide has a unique chemical structure. Capsular PS of serotype 10D was purified from a culture seeded with one colony of strain Cam853 using anion-exchange chromatography. Each fraction was tested for a 10D capsule (in an inhibition enzyme-linked immunosorbent assay [ELISA]) as well as absorbance at 280 and 260 nm for protein and nucleic acid contamination, respectively (Fig. 2A). Fraction 15, which had high concentration of 10D PS with low protein or nucleotide contaminants, was selected for nuclear magnetic resonance (NMR) studies. One-dimensional (1D) ¹H NMR spectra were obtained with Cam853 capsular PS as well as capsular PS from serotypes 6C, 10A, and 39. The latter three serotypes were investigated for direct comparison with Cam853 PS. As shown in Fig. 2B, the 1D ¹H NMR spectra of the anomeric carbons (chemical shifts of 4.6 to 5.6 ppm) are distinct for each serotype, clearly indicating that Cam853 capsule is chemically distinct from the other three. The

TABLE 2 Unique serological profile of strain Cam853 by agglutination^a

Strain	Antiserum				
	Group 10	FS10b	FS10d	FS10f	Type 39
Cam853	–	–	+	–	–
SSISP10A	+	–	+	–	–
SSISP10B	+	+	+	–	–
SSISP10C	+	+	–	+	–
SSISP10F	+	+	–	–	–
SSISP39	–	–	+	–	+

^aCam853 is serologically unique from other serogroup 10 members and serotype 39. Pneumococcal strains SSISP10A, SSISP10B, SSISP10C, SSISP10F, and SSISP39 are from Statens Serum Institut (SSI) (23) and represent serotypes 10A, 10B, 10C, 10F, and 39, respectively. All group-specific, type-specific, and factor-specific (FS) antisera are also from SSI. Cam853 reacted only with FS10d in the agglutination reaction. Negative reactions (–) and positive reactions (+) are shown.

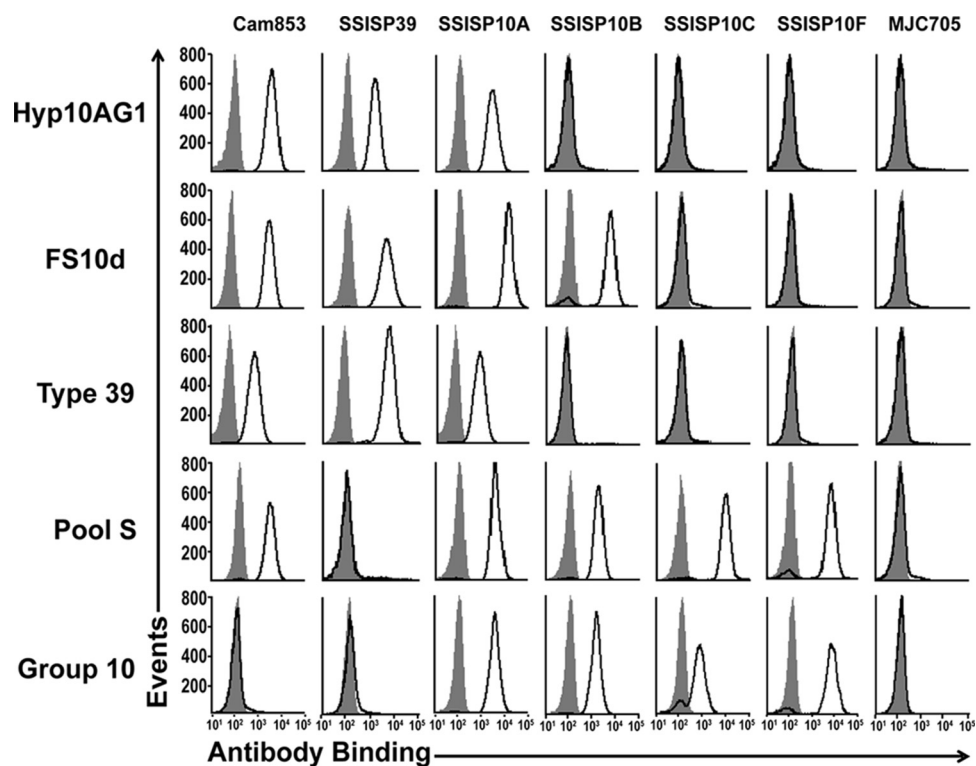


FIG 1 Serological profile of strain Cam853 and other pneumococcal serotypes by flow cytometric analysis. Histograms show fluorescence of a bacterial strain (indicated at the top of each column of histograms) after staining with the indicated serologic reagents (indicated to the left of each row of histograms). Pneumococcal strains SSISP39, SSISP10A, SSISP10B, SSISP10C, SSISP10F, and MJC705 represents serotypes 39, 10A, 10B, 10C, 10F, and 6C, respectively. The solid black lines indicate the fluorescence obtained with primary and secondary antibodies. The gray shaded areas represent control binding (secondary antibody alone). The x axes show the log fluorescence intensity, and the y axes show the number of events (cell counts). Hyp10AG1 is the serotype 10A-specific monoclonal antibody (MAb). All other reagents were obtained from Statens Serum Institut (SSI) (FS10d, factor serum 10d). Each strain was tested three times, and a representative result was shown.

NMR spectrum of 10D showed a minor signal from teichoic acid (3.17 ppm), suggesting that the 10D capsular PS is sufficiently pure for detailed structural studies.

Next, we employed additional NMR techniques to determine the chemical structure of 10D capsular PS (Fig. 3). Shifts in the NMR signals of ^1H and ^{13}C (i.e., chemical shifts) reflect the chemical milieu of each carbon atom, and they can be used to determine the structure of a PS. The ^1H and ^{13}C chemical shifts of 10D PS (Fig. 3A and B and Table 4) were assigned using a set of standard two-dimensional NMR experiments as described in Material and Methods. To facilitate assignments of the chemical shifts for 10D PS, similar NMR experiments were also performed with the 10A, 6C, and 39 PS. Chemical shifts for 10A, 39, and 6C PS were identical to those reported in previous studies (24, 26). The phosphate glycosidic linkage of 10D PS was confirmed by ^1H - ^{31}P correlation (Fig. 3C). Table 4 shows the full assignments for 10D PS, and the deduced chemical structure of 10D PS is shown in Fig. 4 along with those of related serotypes. Taken together, the NMR data unambiguously confirmed that 10D produces a chemically distinct capsule.

Of note, the chemical structure of 10D allowed for the identification of biosynthetic roles of all genes in 10D *cps* locus as summarized in Fig. 4. 10D and 6C share the structure of the two glucose residues at the reducing end as well as two glycosyltransferases, *wchA* and *wciN β* in their *cps* (11, 26). Therefore, both genes are involved in forming the glucose dimer. Ribitol is linked to the glucose dimer, and *wcrO*_{10D} is the only gene encoding a ribitol transferase in 10D *cps*. The remaining structure is formed by the products of genes (*wcrC*, *wcrD*, *wcrF*, and *wcrG*), which are shared between serotypes 10A and 39 as described previously (24).

TABLE 3 Serological profile of strain Cam853 and other related serotypes by flow cytometry^a

Strain	Antiserum							
	Pool E	Pool S	Type 39	Group 10	FS10b	FS10d	FS10f	FS6d
Cam853	+	+	+	-	-	+	-	-
SSISP39	+	-	+	-	-	+	-	-
SSISP10A	+	+	+	+	-	+	-	-
SSISP10B	+	+	-	+	+	+	-	-
SSISP10C	+	+	-	+	+	-	+	-
SSISP10F	+	+	-	+	+	-	-	-
MJC705	-	-	-	-	-	-	-	+

^aPneumococcal strains SSISP10A, SSISP10B, SSISP10C, SSISP10F, and SSISP39 representing serotypes 10A, 10B, 10C, 10F, and 39, respectively, were procured from Statens Serum Institut (SSI), Copenhagen, Denmark. MJC705 strain represents the serotype 6C and was obtained from the Nahm laboratory bacterial strain collection. Cam853 is the 10D *cps*-containing strain and was obtained from during a post-PCV13 introduction colonization study at Angkor Hospital for Children/Cambodia Oxford Medical Research Unit. Pool E and pool S were designed to react with multiple serotypes (58). Group serum reacts to all the serotypes within one serogroup. Type serum reacts only with a single serotype. Factor serum (FS) reacts with a specific epitope in a serotype. All the antisera are polyclonal rabbit antisera and were obtained from SSI. Negative reactions (-) and positive reactions (+) are shown. Cam853 reacted only with FS10d in agglutination reaction and flow cytometry. Cam853 also reacted with pool E and S antisera in flow cytometry, which otherwise showed negative results in agglutination reaction.

Immunization with 10A polysaccharide elicits antibodies opsonizing 10D. Since serologic studies suggested that 10A PS may elicit antibodies cross-reacting with 10D, we investigated whether the 23-valent pneumococcal PS vaccine (PPV23), which contains the 10A serotype, induces antibodies cross-opsonizing serotype 10D. An opsonophagocytosis assay (OPA) with six pairs of pre- and postvaccinated immune sera

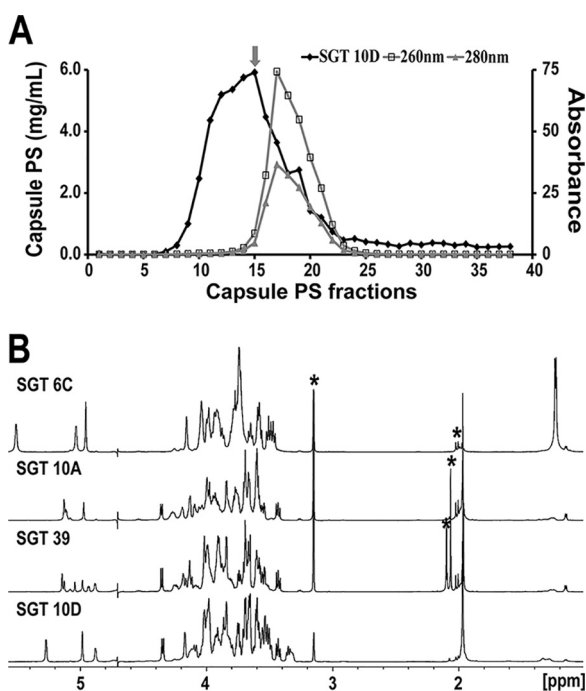


FIG 2 (A) Purification of serotype 10D capsular polysaccharide. Anion-exchange fractions (x axis) were tested for 10D capsule polysaccharide (PS) (solid diamonds), OD₂₆₀ (open squares), and OD₂₈₀ (triangles). 10D polysaccharide was determined by an inhibition ELISA. OD₂₆₀ and OD₂₈₀ reflect nucleic acid and protein contaminations, respectively. Fraction 15 (indicated by an arrow) was selected for nuclear magnetic resonance (NMR) studies. (B) One-dimensional ¹H NMR spectra of a selected region of the ¹H NMR spectra of 6C, 10A, 39, and 10D capsule polysaccharides. Purified capsular polysaccharides of serotypes 39, 10A, and 6C were procured from Statens Serum Institut. Asterisks indicate cell wall polysaccharide-specific peaks. The NMR spectrum of anomeric carbons (4.6 to 5.6 ppm) indicates that 10D capsule polysaccharide is chemically distinct from 39, 10A, and 6C. SGT, serogroup/type.

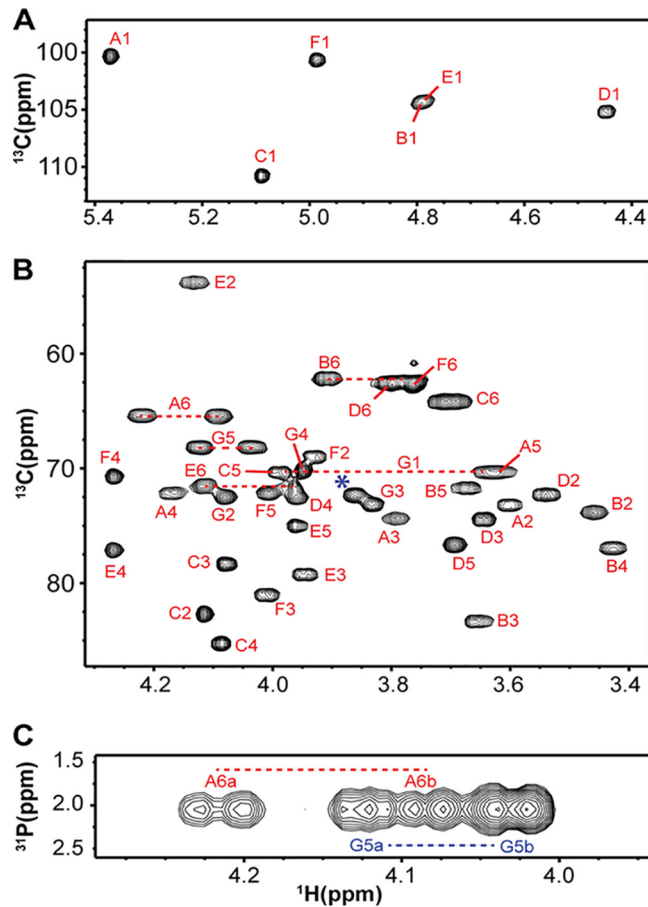


FIG 3 Two-dimensional nuclear magnetic resonance (NMR) spectra of serotype 10D polysaccharide. ^1H - ^{13}C HMQC spectra of anomeric carbons and the other carbons are shown in panels A and B, respectively. (C) ^1H - ^{31}P HMBC spectrum of the ribitol phosphate region. NMR spectra were obtained at 45°C. In the peak labels, the letter indicates the sugar residue (see Table 4), and the number denotes the carbon position on each unit. Peak assignments are shown in Table 4. The asterisk denotes an unassigned peak. Horizontal dashed lines connect the two proton signals associated with the sugar residue at the indicated carbon position.

was performed using serotype 10D (Cam853) and 10A (SSISP10A) strains as targets. Vaccination greatly increased opsonic capacity to both serotypes 10A and 10D ($P < 0.005$ for both), and the mean OPA titer values for 10A and 10D increased about 10-fold following vaccination (Fig. 5A). Furthermore, the increase in OPA titer is specific to 10A and 10D, since preabsorption with homologous 10A PS reduced the 10A or 10D OPA

TABLE 4 ^1H and ^{13}C nuclear magnetic resonance chemical shifts of 10D capsule polysaccharide^a

Residue	Label ^b	NMR chemical shift (proton/carbon shift) (ppm) ^c						
		H1/C1	H2/C2	H3/C3	H4/C4	H5/C5	H6a, H6b/C6	N-Ac
→2- α -D-Glcp-(1→	A	5.37/100.4	3.60/73.2	3.79/74.4	4.17/72.2	3.61/70.39	4.08,4.22/65.5	
→3- α -D-Glcp-(1→	B	4.79/104.6	3.46/73.9	3.65/83.3	3.42/77.1	3.67/71.7	3.77,3.91/62.2	
→ β -D-Galp-(1→	C	5.09/110.8	4.12/82.8	4.08/78.4	4.09/85.4	4.00/70.4	3.70/64.2	
→ β -D-Galp-(1→	D	4.45/105.2	3.54/72.3	3.64/74.4	3.96/72.5	3.69/76.7	3.79,3.80/62.7	
→4- β -D-GalpNAc-(1→	E	4.78/104.1	4.13/53.8	3.95/79.4	4.27/77.1	3.96/75.1	4.11,3.96/71.6	2.07/23.7
→3- α -D-Galp-(1→	F	4.99/100.73	3.93/69.0	4.01/81.1	4.27/70.7	4.00/72.1	3.75,3.76/62.7	
→1-D-Rib-ol-(5→	G	3.64,3.98/70.3	4.08/72.5	3.83/73.3	3.94/70.2	4.03,4.12/68.3		

^aChemical shifts of 10D capsule polysaccharide were recorded at 45°C.

^bEach carbohydrate residue is labeled with a unique letter (A to G).

^cH and C letters indicate proton and carbon, respectively, and the numbers associated with them indicate their respective chemical shift values. The proton and carbon chemical shifts (ppm) are separated by a slash. For each residue, the table shows the chemical shifts of every proton and carbon molecule attached to it at different positions. N-Ac, N-acetylation.

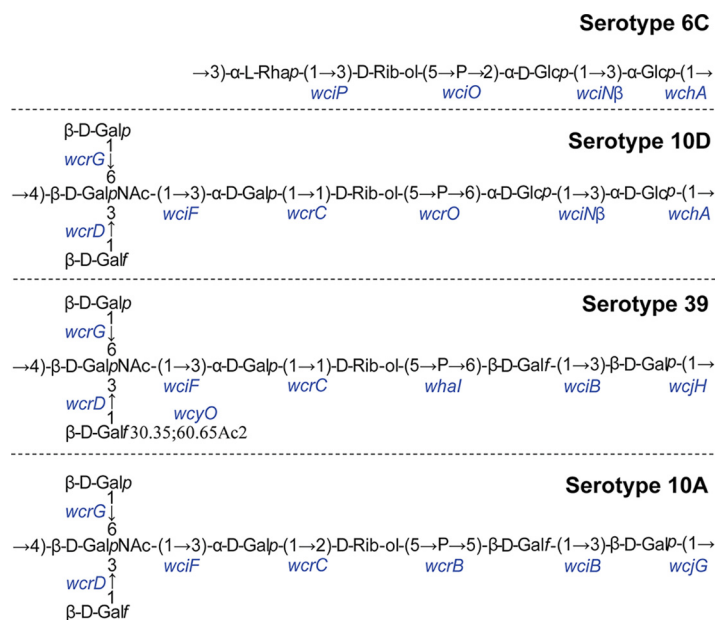


FIG 4 Structure of the serotype 10D capsule polysaccharide in comparison to serotype 6C, 39, and 10A polysaccharide. The structure of the repeating unit of the polysaccharide from each serotype is shown along with the genes (shown in blue) responsible for biosynthesis of the chemical structure. Serotype 10D shares the two glucose residues and their corresponding glycosyltransferase genes with serotype 6C. Ribitol phosphotransferase *wcrO*_{10D} is responsible for the ribitol (5→P→6) glucose linkage. The remaining structure on the left is similar to those in serotypes 39 and 10A, except that the *wcyO* gene mediates O-acetylation in serotype 39, and *wcrC*_{10A} mediates (1→2) linkage in serotype 10A. The structural similarity in Galp, Galf, and GalNac residues in serotypes 10D, 39, and 10A explains the serological cross-reactivity with F510d and Hyp10AG1.

activities almost completely ($P < 0.0001$ for both), but preabsorption with 24F PS (negative control) did not reduce the 10A ($P < 0.0001$) or 10D ($P < 0.001$) OPA activities (Fig. 5B). It is therefore likely that 10A PS would elicit protective antibodies against 10D.

10D cps has a ~6 kb fragment that is syntenic and homologous to a *Streptococcus mitis* cps. To understand how 10D *cps* arose, we searched for the origin of *wcrO*_{10D}, which is absent in the *cps* of serotypes 6C and 39. Phylogenetic analysis of *wcrO*_{10D} identified various genes encoding ribitol phosphate transferases from oral streptococci and pneumococci with a 45 to 60% nucleotide identity (Fig. 6). No genes from pneumococcal *cps* loci showed high homology to *wcrO*_{10D}. However, *wcrO*_{10D} showed high homology (90% nucleotide identity, 94% amino acid identity) to one gene (*RS00925*) from an oral streptococcus strain, SK145 (Fig. 6).

Even more surprising is the fact that SK145 and 10D *cps* displayed very high homology (~94% nucleotide identity) for a very large (~6-kb) *cps* region beginning from the end of the 6C-like region (*wcrO*) to *wzy* (Fig. 7). SK145 *cps* even contained the unique 3' end of *wciNβ*_{10D} and the genetic fragment of *wciNα*_{10D}, both of which came from *wciNα*_{SK145}, including the 7-base overlap that both *wciNα* and *wciNβ* have (bases 6073 to 6079 of 10D *cps*). As a result, *wciNβ*_{10D} is shorter than *wciNβ*_{6C} by 6 nucleotides at the 3' end, and 10D *WciNβ* protein should be shorter than 6C *WciNβ* (373 versus 375 amino acids) (Fig. 7). 5' to the 7-base overlap, 10D *cps* (from *cpsA* to *wciNβ*) is highly homologous (98% nucleotide identity) to the 6C *cps* locus. The 3' end of 10D *cps* (from *wcrG* to *glf*) is more homologous (88% nucleotide identity) to serotype 39 *cps* than to SK145 *cps*. Thus, serotype 10D *cps* might have evolved by double recombination events between serotype 6C, strain SK145, and serotype 39 *cps* loci. However, given that upper respiratory streptococci are genetically very diverse and their *cps* loci have not been extensively studied, it is more likely that there was a single recombination event between serotype 6C and an SK145-like strain harboring all the genes from the *wciNα* genetic fragment to *glf* (~9 kb). The seven bases shared between 6C and SK145 could have facilitated the crossover between serotype 6C *cps* and SK145-like strain (Fig. 7).

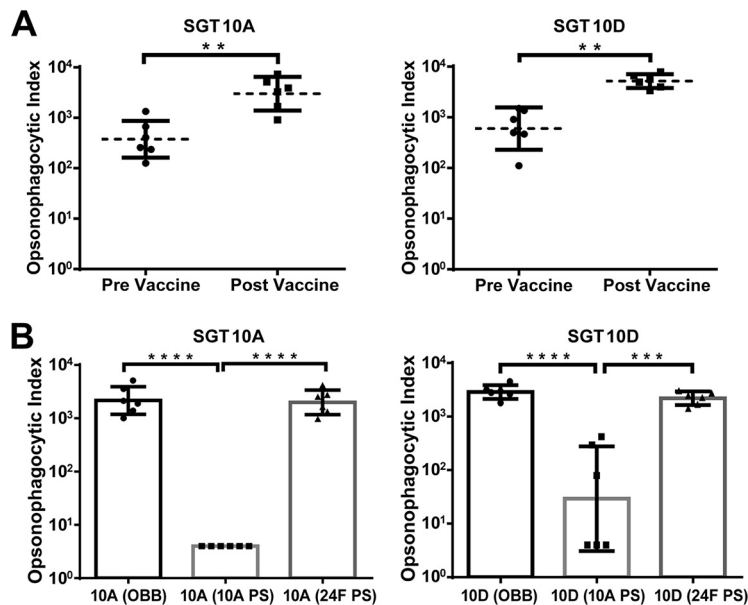


FIG 5 23-valent pneumococcal polysaccharide vaccine (PPV23) elicits opsonic antibodies to serotype 10D. (A) Capacity to opsonize serogroup/type (SGT) 10A (left) or 10D (right) with sera from six adults before and after PPV23 immunization. The dashed black lines represent the mean opsonophagocytic indices. Error bars indicate standard deviations. Data were analyzed by the Mann-Whitney test. (B) Capacity of PPV23 immune sera to opsonize SGT 10A and 10D in the presence of OBB buffer (left), 10A PS (middle), or unrelated PS (right). Inhibitor concentrations were 200 $\mu\text{g}/\text{ml}$ in OBB. Values are means \pm standard deviations (error bars). Data were analyzed by a one-way analysis of variance (ANOVA) with Tukey's multiple-comparison test. Values that are significantly different are indicated by a bar and asterisks as follows: **, $P < 0.005$; ***, $P < 0.001$; ****, $P < 0.0001$. PS, polysaccharide.

DISCUSSION

Pneumococcus is a highly successful pathogen in part due to its high level of capsule diversity, resulting in a plethora of unique serotypes (6). Recent studies show its capsule diversity to be greater than previously appreciated. Although almost a century of studies based on Quellung reactions identified 90 different capsule types (27), the use of MAbs and genetic screens enabled the discovery of nine new capsule types starting with the discovery of serotype 6C in 2007 (11). Herein, we describe the 100th pneumococcal capsule type, named serotype 10D, by providing the complete chemical structure along with identification of biosynthetic roles for all the *cps* genes. In addition, the chemical structure explains its serologic properties: its reaction with polyclonal serum (FS10d) and MAb Hyp10AG1, which targets $\beta\text{Gal}(1-6)$ created by *wcrG* (24).

Phylogenetic analysis of 10D *cps* suggests that 10D *cps* captured the *wcrO*_{10D} gene from an oral streptococcus species, which often contains pneumococcus-like *cps* loci and actually produces capsular PSs (28–32). Specifically, *wcrO*_{10D} sequence has low homology with the four *S. pneumoniae* *wcrO* genes in serotype 33C, 34, 35F, and 36 (40 to 50% amino acid identity) even though pneumococcal *cps* loci have been extensively studied (more than 20,000 pneumococcal isolates [17]). In contrast, *wcrO*_{10D} shows surprisingly high homology to a gene (*RS00925*) in the *cps* gene of *S. mitis* strain SK145. In fact, the *wcrO*_{10D} gene may be common among oral streptococci, since oral streptococcal *cps* loci have not been studied as extensively as pneumococci. Taking all the evidence together, the *wcrO*_{10D} gene has likely originated from oral streptococci, not from pneumococci.

Interspecies capsule gene exchanges have been previously suggested (15). For instance, serotypes 19B and 19C may have emerged when 19F captured an ~ 13 -kb fragment with 10 functional genes from SK564, an *S. mitis* strain (21). Another example is suggested by an ~ 16.5 -kb homologous region with 16 genes shared between *cps* of serotype 2 and strain SK95 (28). Like these previous examples, examination of 10D and

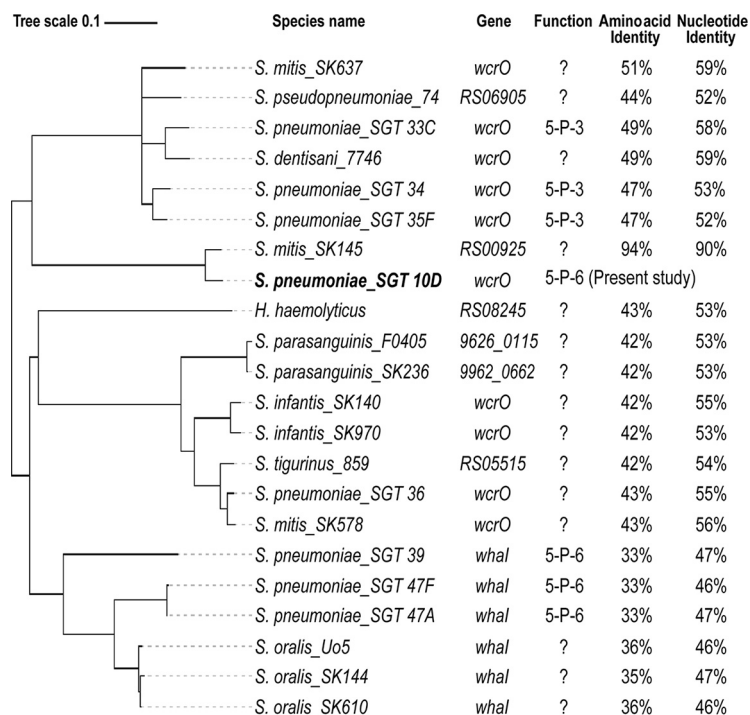


FIG 6 Phylogenetic analysis of the *wcrO*_{10D} gene. The phylogenetic tree shows that the *wcrO*_{10D} gene has the highest homology (90% identical at the nucleotide level and 94% at the amino acid level) with the *S. mitis* SK145 gene RS00925. All the genes from pneumococcal and nonpneumococcal strains depicted in the tree belong to pfam04991 (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). The question mark symbol indicates that the gene function is unknown. The length of the scale bar represents the estimated evolutionary divergence between species based on average genetic distance. A scale bar of 0.1 indicates 10% nucleotide substitution per site. Nucleotide and translated amino acid sequences below 45% and 30% identity, respectively, were eliminated from the analysis. RS06905, RS00925, RS08245, 9626_0115, 9962_0662, and RS05515 are locus tags. The NCBI accession numbers of all the strains used in phylogenetic analysis are provided in Table S1 in the supplemental material. SGT, serogroup/type.

SK145 *cps* also shows a very large (~6-kb-long) syntenic and highly homologous region, which has five functional genes, including *wcrO*_{10D}. Unlike previous examples, however, the 10D *cps* includes at the 5' end of the syntenic region a short nonfunctional gene fragment of *wciN* α that lacks evolutionary constraint for retention. Since SK145 has a full-length transcribable *wciN* α gene, the presence of nonfunctional *wciN* α gene fragment in 10D *cps* suggests recent genetic transfer from strain SK145 to 10D. Moreover, the unique 3' end of *wciN* β _{10D} may be the interspecies recombination crossover point, which was not reported in previous examples (21, 28). Thus, 10D *cps* provides the most compelling evidence for interspecies *cps* exchange with clear directionality.

Interspecies transfer of the *cps* locus would significantly increase the diversity of pneumococcal capsule types since the human upper airway harbors many sources of genes. Oral streptococcus species include not only the mitis group species but also many genetically diverse species that have *cps* loci resembling pneumococcal *cps* (28, 33). Also, genetic materials useful for increasing pneumococcal capsule diversity can be from unrelated Gram-positive or even Gram-negative species. For instance, *Streptococcus thermophilus* has *cps*-like loci (34), and *E. coli* has a lipopolysaccharide (LPS) synthesis locus which resembles pneumococcal *cps* loci (35). While a large genetic survey (involving more than 21,000 strains) in selected sites found only several pneumococcal strains with novel genetic features (17, 21), surveys of additional sites, including the sites with history of extensive usage in PCVs, may reveal additional novel serotypes.

Discoveries of novel capsule types can have a direct impact on PCV usage and

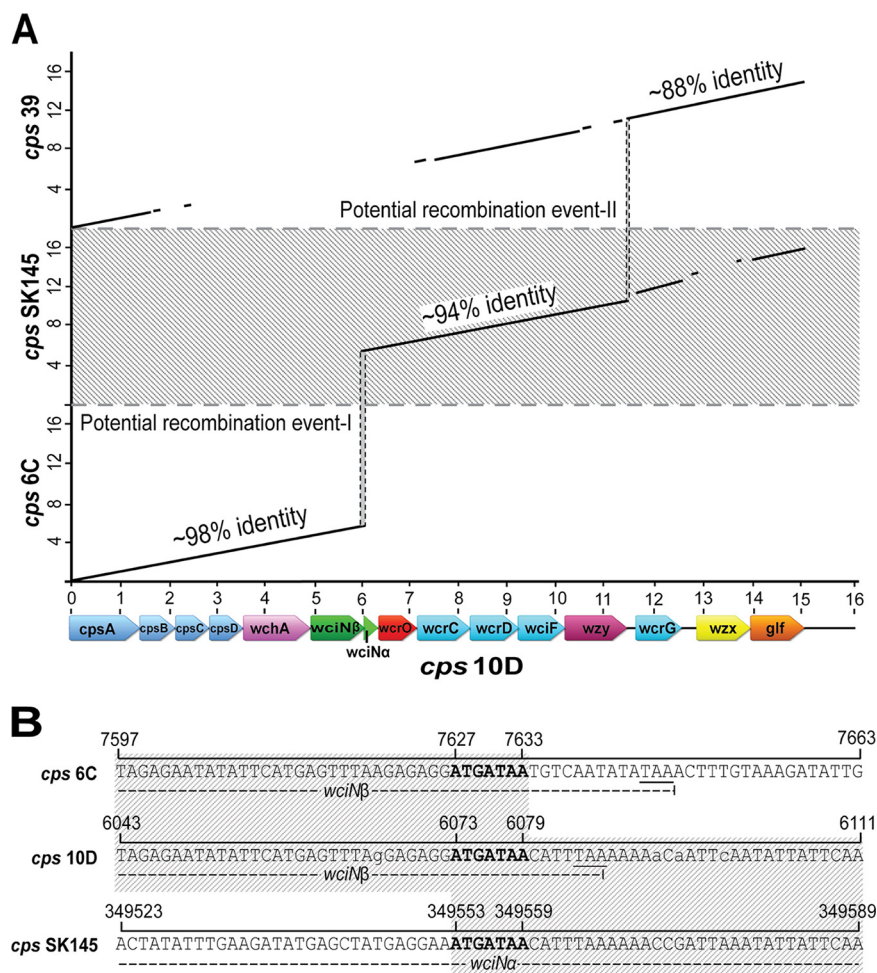


FIG 7 Potential evolutionary origin of serotype 10D capsule biosynthetic loci (*cps*). (A) Comparison of 10D *cps* (SRA accession no. [ERR051587](https://www.ncbi.nlm.nih.gov/SRA/ERR051587)) (21) with 6C *cps* (GenBank accession no. [EF538714](https://www.ncbi.nlm.nih.gov/GenBank/EF538714)), SK145 *cps* ([NZ_JYGS01000001.1](https://www.ncbi.nlm.nih.gov/GenBank/NZ_JYGS01000001.1); contig 1), and 39 *cps* ([CR931711](https://www.ncbi.nlm.nih.gov/SRA/CR931711)) using dot plot analysis. The x axis shows the 10D *cps* gene arrangement; 6C, SK145, and 39 *cps* loci are plotted on the y axis. For strain SK145, reverse complement sequence was used for comparison. The numbers on the x and y axes relate to arbitrary nucleotide base numbers (in kilobases) to show relative size. The solid black lines represent the homologous regions, with the percentage nucleotide identity indicated above the line. The vertical gray-shaded areas indicate overlap regions for the potential recombination events I and II. (B) The nucleotide sequences of *cps* loci for serotype 6C, serotype 10D, and strain SK145 at the transition region for the potential recombination event I. The sequence in bold type indicates the overlap region. Underlined nucleotides indicate the *wciN*_{6C} and *wciN*_{10D} translation stop codons. Lowercase letters in 10D *cps* indicate the nucleotide mismatches. The base numbers indicate the actual nucleotide position in the corresponding *cps* loci. The gray-shaded area indicates the region of homology “upstream” (*cps* 6C and *cps* 10D) and “downstream” (*cps* 10D and *cps* SK145) of the potential recombination site. The dashed horizontal lines indicate the genes involved in the potential recombination event I.

design. Prior to the discovery of serotype 6C, it was mistyped as “6A” (11). Discovery of serotype 6C allowed one to show that 7-valent PCV (PCV7) was cross-protective against 6A but not 6C (12), and helped improve PCV design. Like 6C, epidemiologic surveys have not distinguished 10D from 10A so far and reported increased prevalence of “10A” following the use of PCVs in children (36–39). One should now consider that PCV usage may have increased the prevalence of 10D as well as 10A. New PCVs including 10A PS may provide cross-protection against 10D, since we found 10D-opsonic antibodies in persons vaccinated with PPV23, which contains 10A PS. Since PCV20 under development contains 10A, it would be interesting to examine its effect on 10D IPDs and carriage.

The natural prevalence of different pneumococcal capsule types in the upper respiratory tract represents equilibrium of various selection forces in nature. When a

new selective pressure is introduced, the new force alters the equilibrium demonstrated by serotype replacement following the use of PCVs (7). The selective force could foster interspecies genetic exchange. The appearance of antibiotic resistance genes among pneumococci followed the extensive use of antibiotics (40–43). An intriguing question is whether the use of PCVs increases capturing of oral streptococcus *cps* by pneumococcus resulting in “vaccine-escape” strains (31). Another interesting possibility is that PCVs may impose selective pressure onto oral streptococci as well (28–30) and select for nonvaccine capsule types, since they appear to express capsule types that are similar to PCV serotypes. In summary, the large pneumococcal capsule diversity continues to be a significant threat to human health, and the unusual sources of capsule diversity are contributing to the problem. Thus, the long-term strategy of using PCVs must include improved serotype knowledge of pneumococci as well as oral streptococci residing in our upper respiratory tract.

MATERIALS AND METHODS

Bacterial strains and cultivation. Strains Cam657 (GPS_KH_COMRU657; SRA accession no. ERR2681221) and Cam853 (GPS_KH_COMRU853; ERR2680924) were collected between 2016 and 2017 during a post-PCV13 introduction colonization study at the Angkor Hospital for Children/Cambodia Oxford Medical Research Unit (20). Strain PATH4346 (GPS_US_PATH4346; ERR1453813) was collected in 2007 as a part of cross-sectional population-based study in Ethiopian children during mass azithromycin distribution for trachoma (19). As a part of the Global Pneumococcal Sequencing project (<https://www.pneumogen.net/gps/>), the whole genomes of these carriage strains were sequenced and these strains were genetically typed as 39X (17).

Pneumococcal strains SSISP10A, SSISP10B, SSISP10C, SSISP10F, and SSISP39, representing serotypes 10A, 10B, 10C, 10F, and 39 respectively, were from Statens Serum Institut (SSI) (Copenhagen, Denmark); a serotype 6C strain (MJC705) was from the Nahm laboratory bacterial strain collection.

All pneumococcal strains were isolated on blood agar plates with 5% sheep blood (Remel Laboratories, Lenexa, KS) after overnight incubation at 37°C with 5% CO₂. Isolated colonies were inoculated into Todd-Hewitt broth with 5% yeast extract (THY) and grown to mid-log density (optical density at 600 nm [OD₆₀₀] of 0.5), and bacteria were frozen in THY with 15% glycerol at –80°C for later use. All isolates had colony morphology typical of pneumococci and were identified as pneumococcus by matrix-assisted laser desorption/ionization–time of flight mass spectrometry (MALDI-TOF MS) using Vitek MS Knowledge Base V3.2 (bioMérieux) (44).

Genetic confirmation of 39X (serotype 10D) genotype. The 39X genotype of strain Cam853 was confirmed using the primers 5943 (5'-AGTTAGGTAGTGCTGCTTCC-3') and 3773 (5'-CTATCAGCAGTAGT TAGTGTAATTATC-3') to amplify a 1,641-bp unique region (*wciNβ-wcrO-wcrC*) of the 39X *cps* locus. The PCR mixture was 25 μl of 1× *Ex Taq* PCR buffer containing 50 to 100 ng of genomic DNA, 0.3 U of *Ex Taq* polymerase (3 U/μl; TaKaRa), 1 μl deoxynucleotide triphosphates (dNTPs) (2.5 mM each [Fermentas, USA]), and 1 μl of each forward and reverse primer (each 10 μM). Thermal cycling was performed in GeneAmp PCR system 9700 (Applied Biosystems, USA) under the following conditions: 94°C for 10 min, followed by 35 amplification cycles, with 1 cycle consisting of 94°C for 30 s, 58°C for 30 s, and 72°C for 1 min, and a final extension step at 72°C for 10 min. The PCR products were separated by electrophoresis on 1.5% agarose gel. The amplified products were purified using a QIA quick PCR purification kit (Qiagen, Germany) following the manufacturer's protocol. Purified PCR products were subjected to bidirectional sequencing which was performed by Heflin Center Genomics Core Lab at the University of Alabama at Birmingham (UAB).

Serological analysis of serotype 10D isolates. The serological properties of the serotype 10D (39X) isolates were assessed by the slide agglutination reaction and flow cytometry as previously described (27, 45) using a panel of polyclonal antisera obtained from the SSI (Copenhagen, Denmark) as detailed in Table 3, or using our in-house 10A- and 6C/6D-specific monoclonal antibodies (MAbs) as depicted in Fig. 1 and Fig. S4 in the supplemental material (46). The MAbs were produced as described previously (23). For flow cytometry, frozen bacterial stocks were thawed and washed in fluorescence-activated cell sorting (FACS) buffer (1× phosphate-buffered saline [PBS], 3% fetal bovine serum [FBS], 0.1% NaN₃), incubated with 1:1,000 dilutions of polyclonal antisera, or 1:100 and 1:200 dilutions of 10A-specific and 6C/6D-specific MAbs, respectively. After the isolates were washed, they were labeled with a phycoerythrin (PE)-conjugated anti-rabbit Ig secondary antibody (Southern Biotech, Birmingham, AL) at 1:1,000 dilution or a PE-conjugated anti-mouse Ig secondary antibody (BD Biosciences, San Jose, CA) at 1:200 dilution. After the bacteria were washed, their fluorescence was measured with a BD Accuri flow cytometer. Data were analyzed in FCS Express version 3.0. The assay was performed with the indicated bacterial strains and antisera three times.

Inhibition ELISA. To detect the capsule polysaccharide (PS) during purification, an inhibition-type enzyme-linked immunosorbent assay (ELISA) was performed. Briefly, the wells of ELISA plates (Corning Costar Corp., Acton, MA) were coated at 37°C with 1 μg/ml of 10A capsular PS (SSI, Copenhagen, Denmark) for 5 h in phosphate-buffered saline. After the plates were washed with PBS containing 0.05% Tween 20, 50-μl portions of the samples containing PS were added to the wells along with 50 μl of factor serum 10d (FS10d) at 1:10,000 dilutions. After 1 h of incubation at room temperature, the plates were washed and incubated for 1 h with 100 μl of alkaline phosphatase-conjugated goat anti-rabbit Ig

(Southern Biotech) at 1:3,000 dilutions. The amount of the enzyme immobilized to the wells was then determined by incubating paranitrophenyl phosphate substrate (Sigma) in diethanolamine buffer for 1 to 2 h at room temperature and measuring the OD₄₀₅ with a microplate reader (BioTek Instruments Inc., Winooski, VT). Purified 10A PS (SSI, Copenhagen, Denmark) was used as an assay standard (Fig. S5).

Purification of capsular polysaccharide. A single colony of the Cam853 strain was grown in 10 ml of THY broth and then expanded in 1 liter of a chemically defined medium (47) supplemented with choline chloride (1 g/liter), sodium bicarbonate (2.5 g/liter), and cysteine HCl (0.73 g/liter). After overnight incubation at 37°C, bacteria were separated from the supernatant by centrifugation (15,344 × *g*, 30 min, 4°C), washed, and resuspended in 18 ml of 0.9% NaCl. After the pH was adjusted to 7, the bacterial suspension was added with 100 μl of 10% sodium deoxycholate, 200 μl of mutanolysin (10 U/μl), and 1.7 ml of lysozyme (11.8 mg/ml) and incubated at 37°C for 48 to 72 h. The resulting lysate was centrifuged to remove debris, dialyzed overnight against 4 liters of 5 mM Tris (pH 7.3) (dialysis tubing, 3,500-molecular-weight cutoff). The dialyzed product was applied to a DEAE Sepharose (GE Healthcare, Uppsala, Sweden) anion-exchange column (40 ml). After the column was washed with 2.5 volumes of 5 mM Tris (pH 7.4), elution was performed with a linear gradient of NaCl ranging from 0 M to 0.4 M over 37 fractions (2.5 ml/fraction). Each fraction was tested for OD₂₆₀, OD₂₈₀, and 10D PS with an inhibition ELISA as described above. Fraction 15 was selected for PS studies and was lyophilized after dialysis.

Opsonophagocytosis assay. To investigate whether the 10A serotype elicits cross-opsonizing antibodies to serotype 10D, we adapted a well-characterized UAB opsonophagocytosis assay (OPA) (48) (and described in detail at <https://www.vaccine.uab.edu>). OPA was performed with six pairs of pre- and post-PPV23-vaccinated immune sera using serogroup/type (SGT) 10A and 10D as targets. Briefly, 30 μl of bacteria suspended in OBB (Hanks' buffer supplemented with 0.1% gelatin and 5% fetal calf serum) was mixed with 10 μl of baby rabbit serum (BRS) of specified concentration, and 40 μl of differentiated HL60 cells (10⁷ cells/ml) in OBB. The mixture was incubated with shaking (700 rpm) for 45 min at 37°C with 5% CO₂. Ten microliters from each well was spotted onto THY agar plates, and the bacterial colonies were counted after overnight incubation. To examine the specificity of the 10A-induced immune response, postvaccinated sera were preabsorbed with the homologous 10A capsule PS (200 μg/ml in OBB) or heterologous 24F (negative-control) capsule PS (200 μg/ml in OBB), incubated for 30 min at room temperature, followed by the typical OPA procedure. As a control, each sample was tested without the preabsorption of any external capsule PS.

NMR spectroscopy. ¹H-¹H and ¹H-¹³C nuclear magnetic resonance (NMR) data were collected at 45°C on a Bruker Avance II (700 MHz ¹H) or Avance III (600 MHz ¹H) spectrometers equipped with cryogenic triple-resonance probes, processed with NMRPIPE (49) and analyzed with NMRVIEW (50). NMR samples were prepared by dissolving ~5 mg of lyophilized PS in 0.5 ml of 99.99% D₂O (Cambridge Isotope Laboratories). Complete assignments of ¹H and ¹³C signals were achieved by two-dimensional nuclear Overhauser spectroscopy (¹H-¹H NOESY), correlation spectroscopy (¹H-¹H COSY), total correlation spectroscopy (¹H-¹H TOCSY), heteronuclear multiple quantum coherence (¹H-¹³C HMQC), and heteronuclear multiple bond correlation (¹H-¹³C HMBC) spectra. HDO signal was used as a reference. For the ¹H-³¹P HMBC experiment, ~5 mg lyophilized PS was dissolved in 200 μl of 99.99% D₂O (Cambridge Isotope Laboratories) and transferred into a 3-mm NMR tube. NMR spectra were recorded at 45°C on a Varian Inova DD2 NMR spectrometer (¹H, 599.7 MHz) equipped with a room-temperature 3-mm inverse broadband probe (Complex Carbohydrate Research Center, University of Georgia, Athens, GA). The ¹H-³¹P HMBC spectrum was collected with 2,000 points, 64 increments, four scans per increment, and spectral widths of 7,184 Hz (¹H) and 4,856 Hz (³¹P). The HMBC long-range transfer delay corresponded to a coupling of 8 Hz. ³¹P chemical shifts were referenced to 85% H₃PO₄ at 0 ppm.

Genomic analysis. To predict the origin and function of the *wcrO*_{10D} gene, we searched for the homologous genes in oral streptococcal and pneumococcal sequence data in GenBank. Phylogenetic analyses were conducted based on genetic distance using Geneious prime v2019.2 (Biomatters). Full-length nucleotide and translated amino acid sequences above 45% and 30% identity, respectively (multiple alignments by MAFFT plug-in) were used to construct the phylogenetic tree based on the neighbor-joining method using Tamura-Nei genetic distance model with 1,000 bootstrap replicates. Display and manipulation of the phylogenetic tree (Newick format) were performed using the online tool Interactive Tree of Life (51). Gene name and function were predicted based on the already published functions and linkage specificities of the pneumococcal *cps* genes (52).

To infer the probable evolutionary origin of serotype 10D, we examined the 10D *cps* locus (SRA accession no. [ERR051587](https://www.ncbi.nlm.nih.gov/sra/ERR051587)) (21) for synteny and homology against serotype 6C (GenBank accession no. [EF538714](https://www.ncbi.nlm.nih.gov/nuccore/EF538714)), strain SK145 ([NZ_JYGS01000001.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JYGS01000001.1); contig 1), and serotype 39 *cps* loci ([CR931711](https://www.ncbi.nlm.nih.gov/nuccore/CR931711)) by dot plot analysis using YASS (53). The SK145 GenBank sequence was in the opposite orientation, and its reverse complement sequence was used for comparison. Forward strands of each *cps* locus were analyzed with the following parameters: scoring matrix match, +1; transversion, -5 (no correction); gap costs (opening -16 and extension -4), E-value threshold = 10, and X-drop threshold = 30. Recombination crossover regions were identified from a dot plot using Geneious prime.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 0.8 MB.

FIG S2, TIF file, 0.7 MB.

FIG S3, TIF file, 0.6 MB.

FIG S4, TIF file, 0.7 MB.

FIG S5, TIF file, 0.6 MB.

TABLE S1, DOCX file, 0.01 MB.

ACKNOWLEDGMENTS

This work was supported with funding from the National Institutes of Health (AG-050607) to M.H.N. Isolates were collected through the GPS project funded by the Bill and Melinda Gates Foundation (OPP1034556). The High-Field NMR facility at the University of Alabama at Birmingham (UAB) was established through the NIH (1S10RR026478) and is currently supported by the comprehensive cancer center (NCI grant P30 CA013148). The NMR facility at the University of Georgia is supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, U.S. Department of Energy (grant DE-SC0015662 to Parastoo Azadi at the Complex Carbohydrate Research Center).

Author contributions: F.G., designed and performed all the experiments, data analysis and interpretation, drafting and critical revision of the manuscript; J.S.S., NMR experimentation and analysis, critical revision of the manuscript; L.M., critical revision of the manuscript; A.J.V.T., critical revision of the manuscript; S.D.B., critical revision of the manuscript; S.W.L., critical revision of the manuscript; R.A.G., critical revision of the manuscript; P.T., provision of materials, critical revision of the manuscript; J.D.K., critical revision of the manuscript; R.F.B., critical revision of the manuscript; M.H.N., supervised the group, devised and executed the project, had full access to all study data, takes responsibility for the integrity and accuracy of the data analysis, drafting and critical revision of the manuscript.

UAB has intellectual property rights on some reagents used in the study. F.G., J.S.S., and M.H.N. are UAB employees. We declare that we have no other relevant conflicts of interest.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

- Pasteur L. 1881. Note sur la maladie nouvelle provoquée par la salive d'un enfant mort de la rage. *Bull Acad Méd (Paris)* 10:94–103.
- GBD 2015 LRI Collaborators. 2017. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Infect Dis* 17:1133–1161. [https://doi.org/10.1016/S1473-3099\(17\)30396-1](https://doi.org/10.1016/S1473-3099(17)30396-1).
- Avery OT, Dubos R. 1931. The protective action of a specific enzyme against type III pneumococcus infection in mice. *J Exp Med* 54:73–89. <https://doi.org/10.1084/jem.54.1.73>.
- Avery OT, Macleod CM, McCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exp Med* 79:137–158. <https://doi.org/10.1084/jem.79.2.137>.
- Andrews NJ, Waight PA, Burbidge P, Pearce E, Roalfe L, Zancolli M, Slack M, Ladhani SN, Miller E, Goldblatt D. 2014. Serotype-specific effectiveness and correlates of protection for the 13-valent pneumococcal conjugate vaccine: a postlicensure indirect cohort study. *Lancet Infect Dis* 14:839–846. [https://doi.org/10.1016/S1473-3099\(14\)70822-9](https://doi.org/10.1016/S1473-3099(14)70822-9).
- Geno KA, Gilbert GL, Song JY, Skovsted IC, Klugman KP, Jones C, Konradsen HB, Nahm MH. 2015. Pneumococcal capsules and their types: past, present, and future. *Clin Microbiol Rev* 28:871–899. <https://doi.org/10.1128/CMR.00024-15>.
- Hausdorff WP, Hanage WP. 2016. Interim results of an ecological experiment — conjugate vaccination against the pneumococcus and serotype replacement. *Hum Vaccin Immunother* 12:358–374. <https://doi.org/10.1080/21645515.2015.1118593>.
- Weinberger DM, Malley R, Lipsitch M. 2011. Serotype replacement in disease after pneumococcal vaccination. *Lancet* 378:1962–1973. [https://doi.org/10.1016/S0140-6736\(10\)62225-8](https://doi.org/10.1016/S0140-6736(10)62225-8).
- Grabenstein JD, Klugman KP. 2012. A century of pneumococcal vaccination research in humans. *Clin Microbiol Infect* 18(Suppl 5):15–24. <https://doi.org/10.1111/j.1469-0691.2012.03943.x>.
- Mulholland K, Satzke C. 2012. Serotype replacement after pneumococcal vaccination. *Lancet* 379:1387. (Letter.) [https://doi.org/10.1016/S0140-6736\(12\)60588-1](https://doi.org/10.1016/S0140-6736(12)60588-1). (Reply, 379:1388–1389, [https://doi.org/10.1016/S0140-6736\(12\)60590-X](https://doi.org/10.1016/S0140-6736(12)60590-X).)
- Park IH, Pritchard DG, Cartee R, Brandao A, Brandileone MC, Nahm MH. 2007. Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* 45:1225–1233. <https://doi.org/10.1128/JCM.02199-06>.
- Nahm MH, Lin J, Finkelstein JA, Pelton SI. 2009. Increase in the prevalence of the newly discovered pneumococcal serotype 6C in the nasopharynx after introduction of pneumococcal conjugate vaccine. *J Infect Dis* 199:320–325. <https://doi.org/10.1086/596064>.
- Gertz RE, Jr, Li Z, Pimenta FC, Jackson D, Juni BA, Lynfield R, Jorgensen JH, Carvalho MDG, Beall BW, Active Bacterial Core Surveillance Team. 2010. Increased penicillin nonsusceptibility of nonvaccine-serotype invasive pneumococci other than serotypes 19A and 6A in post-7-valent conjugate vaccine era. *J Infect Dis* 201:770–775. <https://doi.org/10.1086/650496>.
- Beall BW, Gertz RE, Hulkower RL, Whitney CG, Moore MR, Brueggemann AB. 2011. Shifting genetic structure of invasive serotype 19A pneumococci in the United States. *J Infect Dis* 203:1360–1368. <https://doi.org/10.1093/infdis/jir052>.
- Wyres KL, Lambertsen LM, Croucher NJ, McGee L, von Gottberg A, Linares J, Jacobs MR, Kristinsson KG, Beall BW, Klugman KP, Parkhill J, Hakenbeck R, Bentley SD, Brueggemann AB. 2013. Pneumococcal capsular switching: a historical perspective. *J Infect Dis* 207:439–449. <https://doi.org/10.1093/infdis/jis703>.
- Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, McGee L, von Gottberg A, Song JH, Ko KS, Pichon B, Baker S, Parry CM, Lambertsen LM, Shahinas D, Pillai DR, Mitchell TJ, Dougan G, Tomasz A,

- Klugman KP, Parkhill J, Hanage WP, Bentley SD. 2011. Rapid pneumococcal evolution in response to clinical interventions. *Science* 331: 430–434. <https://doi.org/10.1126/science.1198545>.
17. van Tonder AJ, Gladstone RA, Lo SW, Nahm MH, Du Plessis M, Cornick J, Kwambana-Adams B, Madhi SA, Hawkins PA, Benisty R, Dagan R, Everett D, Antonio M, Klugman KP, von Gottberg A, Breiman RF, McGee L, Bentley SD, The Global Pneumococcal Sequencing Consortium. 2019. Putative novel cps loci in a large global collection of pneumococci. *Microb Genom* 5:e000274. <https://doi.org/10.1099/mgen.0.000274>.
 18. Chewapreecha C, Harris SR, Croucher NJ, Turner C, Marttinen P, Cheng L, Pessia A, Aanensen DM, Mather AE, Page AJ, Salter SJ, Harris D, Nosten F, Goldblatt D, Corander J, Parkhill J, Turner P, Bentley SD. 2014. Dense genomic sampling identifies highways of pneumococcal recombination. *Nat Genet* 46:305–309. <https://doi.org/10.1038/ng.2895>.
 19. Keenan JD, Sahl I, McGee L, Cevallos V, Vidal JE, Chochua S, Hawkins P, Gebre T, Tadesse Z, Emerson PM, Gaynor BD, Lietman TM, Klugman KP. 2016. Nasopharyngeal pneumococcal serotypes before and after mass azithromycin distributions for trachoma. *J Pediatric Infect Dis Soc* 5:222–226. <https://doi.org/10.1093/jpids/piu143>.
 20. Turner P, Leab P, Ly S, Sao S, Miliya T, Heffelfinger JD, Batmunkh N, Lessa FC, Walldorf JA, Hyde TB, Ork V, Hossain MS, Gould KA, Hinds J, Cooper BS, Ngoun C, Turner C, Day N. 2020. Impact of 13-valent pneumococcal conjugate vaccine on colonization and invasive disease in Cambodian children. *Clin Infect Dis* 70:1580–1588. <https://doi.org/10.1093/cid/ciz481>.
 21. Mostowy RJ, Croucher NJ, De Maio N, Chewapreecha C, Salter SJ, Turner P, Aanensen DM, Bentley SD, Didelot X, Fraser C. 2017. Pneumococcal capsule synthesis locus *cps* as evolutionary hotspot with potential to generate novel serotypes by recombination. *Mol Biol Evol* 34:2537–2554. <https://doi.org/10.1093/molbev/msx173>.
 22. Burton RL, Geno KA, Saad JS, Nahm MH. 2016. Pneumococcus with the “6E” *cps* locus produces serotype 6B capsular polysaccharide. *J Clin Microbiol* 54:967–971. <https://doi.org/10.1128/JCM.03194-15>.
 23. Yu J, Lin J, Kim KH, Benjamin WH, Jr, Nahm MH. 2011. Development of an automated and multiplexed serotyping assay for *Streptococcus pneumoniae*. *Clin Vaccine Immunol* 18:1900–1907. <https://doi.org/10.1128/CVI.05312-11>.
 24. Bush CA, Yang J, Yu B, Cisar JO. 2014. Chemical structures of *Streptococcus pneumoniae* capsular polysaccharide type 39 (CPS39), CPS47F, and CPS34 characterized by nuclear magnetic resonance spectroscopy and their relation to CPS10A. *J Bacteriol* 196:3271–3278. <https://doi.org/10.1128/JB.01731-14>.
 25. Thompson A, Lamberth E, Severs J, Scully I, Tarabar S, Ginis J, Jansen KU, Gruber WC, Scott DA, Watson W. 2019. Phase 1 trial of a 20-valent pneumococcal conjugate vaccine in healthy adults. *Vaccine* 37: 6201–6207. <https://doi.org/10.1016/j.vaccine.2019.08.048>.
 26. Cai P, Moran J, Pavliak V, Deng C, Khoury N, Marcq O, Ruppen ME. 2012. NMR structural analysis of the capsular polysaccharide from *Streptococcus pneumoniae* serotype 6C. *Carbohydr Res* 351:98–107. <https://doi.org/10.1016/j.carres.2012.01.017>.
 27. Henrichsen J. 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* 33:2759–2762. <https://doi.org/10.1128/JCM.33.10.2759-2762.1995>.
 28. Skov Sorensen UB, Yao K, Yang Y, Tettelin H, Kilian M. 2016. Capsular polysaccharide expression in commensal *Streptococcus* species: genetic and antigenic similarities to *Streptococcus pneumoniae*. *mBio* 7:e01844–16. <https://doi.org/10.1128/mBio.01844-16>.
 29. Pimenta F, Gertz RE, Jr, Park SH, Kim E, Moura I, Milucky J, Roupheal N, Farley MM, Harrison LH, Bennett NM, Bigogo G, Feikin DR, Breiman R, Lessa FC, Whitney CG, Rajam G, Schiffer J, Carvalho MDG, Beall B. 2018. *Streptococcus infantis*, *Streptococcus mitis*, and *Streptococcus oralis* strains with highly similar *cps5* loci and antigenic relatedness to serotype 5 pneumococci. *Front Microbiol* 9:3199. <https://doi.org/10.3389/fmicb.2018.03199>.
 30. Lessa FC, Milucky J, Roupheal NG, Bennett NM, Talbot HK, Harrison LH, Farley MM, Walston J, Pimenta F, Gertz RE, Rajam G, Carvalho MG, Beall B, Whitney CG. 2018. *Streptococcus mitis* expressing pneumococcal serotype 1 capsule. *Sci Rep* 8:17959. <https://doi.org/10.1038/s41598-018-35921-3>.
 31. Nahm MH, Brissac T, Kilian M, Vlach J, Orihuela CJ, Saad JS, Ganaie F. 7 September 2019. Pneumococci can become virulent by acquiring a new capsule from oral streptococci. *J Infect Dis* <https://doi.org/10.1093/infdis/jiz456>.
 32. Beall B. 7 September 2019. Potential epidemiologic and historical implications of capsular serotypes shared by pneumococci and their non-pneumococcal relatives. *J Infect Dis* <https://doi.org/10.1093/infdis/jiz457>.
 33. Kilian M, Poulsen K, Blomqvist T, Hävarstein LS, Bek-Thomsen M, Tettelin H, Sørensen UBS. 2008. Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS One* 3:e2683. <https://doi.org/10.1371/journal.pone.0002683>.
 34. Broadbent JR, McMahon DJ, Welker DL, Oberg CJ, Moineau S. 2003. Biochemistry, genetics, and applications of exopolysaccharide production in *Streptococcus thermophilus*: a review. *J Dairy Sci* 86:407–423. [https://doi.org/10.3168/jds.S0022-0302\(03\)73619-4](https://doi.org/10.3168/jds.S0022-0302(03)73619-4).
 35. Schnaitman CA, Klena JD. 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol Rev* 57:655–682. <https://doi.org/10.1128/MMBR.57.3.655-682.1993>.
 36. Ben-Shimol S, Givon-Lavi N, Greenberg D, Dagan R. 2016. Pneumococcal nasopharyngeal carriage in children <5 years of age visiting the pediatric emergency room in relation to PCV7 and PCV13 introduction in southern Israel. *Hum Vaccin Immunother* 12:268–276. <https://doi.org/10.1080/21645515.2015.1095414>.
 37. Weinberger R, von Kries R, van der Linden M, Rieck T, Siedler A, Falkenhorst G. 2018. Invasive pneumococcal disease in children under 16 years of age: incomplete rebound in incidence after the maximum effect of PCV13 in 2012/13 in Germany. *Vaccine* 36:572–577. <https://doi.org/10.1016/j.vaccine.2017.11.085>.
 38. Levy C, Ouldali N, Caeyaex L, Angoulvant F, Varon E, Cohen R. 2019. Diversity of serotype replacement after pneumococcal conjugate vaccine implementation in Europe. *J Pediatr* 213:252–253.e3. <https://doi.org/10.1016/j.jpeds.2019.07.057>.
 39. Kandasamy R, Voysey M, Collins S, Berbers G, Robinson H, Noel I, Hughes H, Ndimah S, Gould K, Fry N, Sheppard C, Ladhani S, Snape MD, Hinds J, Pollard AJ. 2020. Persistent circulation of vaccine serotypes and serotype replacement after 5 years of infant immunization with 13-valent pneumococcal conjugate vaccine in the United Kingdom. *J Infect Dis* 221: 1361–1370. <https://doi.org/10.1093/infdis/jiz178>.
 40. Dowson CG, Coffey TJ, Kell C, Whitley RA. 1993. Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of *Streptococcus mitis* in the formation of a low affinity PBP2B in *S. pneumoniae*. *Mol Microbiol* 9:635–643. <https://doi.org/10.1111/j.1365-2958.1993.tb01723.x>.
 41. Sibold C, Henrichsen J, König A, Martin C, Chalkley L, Hakenbeck R. 1994. Mosaic *pbpX* genes of major clones of penicillin-resistant *Streptococcus pneumoniae* have evolved from *pbpX* genes of a penicillin-sensitive *Streptococcus oralis*. *Mol Microbiol* 12:1013–1023. <https://doi.org/10.1111/j.1365-2958.1994.tb01089.x>.
 42. Janoir C, Podglajen I, Kitzis MD, Poyart C, Gutmann L. 1999. In vitro exchange of fluoroquinolone resistance determinants between *Streptococcus pneumoniae* and viridans streptococci and genomic organization of the *parE-parC* region in *S. mitis*. *J Infect Dis* 180:555–558. <https://doi.org/10.1086/314888>.
 43. Cerdá Zolezzi P, Laplana LM, Calvo CR, Cepero PG, Erazo MC, Gómez-Lus R. 2004. Molecular basis of resistance to macrolides and other antibiotics in commensal viridans group streptococci and *Gemella* spp. and transfer of resistance genes to *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 48:3462–3467. <https://doi.org/10.1128/AAC.48.9.3462-3467.2004>.
 44. Porte L, Garcia P, Braun S, Ulloa MT, Lafourcade M, Montana A, Miranda C, Acosta-Jamett G, Weitzel T. 2017. Head-to-head comparison of Microflex LT and Vitek MS systems for routine identification of microorganisms by MALDI-TOF mass spectrometry in Chile. *PLoS One* 12:e0177929. <https://doi.org/10.1371/journal.pone.0177929>.
 45. Geno KA, Saad JS, Nahm MH. 2017. Discovery of novel pneumococcal serotype 35D, a natural wciG-deficient variant of serotype 35B. *J Clin Microbiol* 55:1416–1425. <https://doi.org/10.1128/JCM.00054-17>.
 46. Oliver MB, van der Linden MP, Kuntzel SA, Saad JS, Nahm MH. 2013. Discovery of *Streptococcus pneumoniae* serotype 6 variants with glycosyltransferases synthesizing two differing repeating units. *J Biol Chem* 288:25976–25985. <https://doi.org/10.1074/jbc.M113.480152>.
 47. van de Rijn I, Kessler RE. 1980. Growth characteristics of group A streptococci in a new chemically defined medium. *Infect Immun* 27: 444–448. <https://doi.org/10.1128/IAI.27.2.444-448.1980>.
 48. Burton RL, Nahm MH. 2006. Development and validation of a fourfold multiplexed opsonization assay (MOPA4) for pneumococcal antibodies. *Clin Vaccine Immunol* 13:1004–1009. <https://doi.org/10.1128/CVI.00112-06>.
 49. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe:

- a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293. <https://doi.org/10.1007/bf00197809>.
50. Johnson BA, Blevins RA. 1994. NMR View: a computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4:603–614. <https://doi.org/10.1007/BF00404272>.
 51. Letunic I, Bork P. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* 39:W475–W478. <https://doi.org/10.1093/nar/gkr201>.
 52. Aanensen DM, Mavroidi A, Bentley SD, Reeves PR, Spratt BG. 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. *J Bacteriol* 189:7856–7876. <https://doi.org/10.1128/JB.00837-07>.
 53. Noe L, Kucherov G. 2005. YASS: enhancing the sensitivity of DNA similarity search. *Nucleic Acids Res* 33:W540–W543. <https://doi.org/10.1093/nar/gki478>.
 54. Calix JJ, Nahm MH. 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. *J Infect Dis* 202:29–38. <https://doi.org/10.1086/653123>.
 55. Calix JJ, Porambo RJ, Brady AM, Larson TR, Yother J, Abeygunwardana C, Nahm MH. 2012. Biochemical, genetic, and serological characterization of two capsule subtypes among *Streptococcus pneumoniae* serotype 20 strains: discovery of a new pneumococcal serotype. *J Biol Chem* 287:27885–27894. <https://doi.org/10.1074/jbc.M112.380451>.
 56. Park IH, Geno KA, Yu J, Oliver MB, Kim KH, Nahm MH. 2015. Genetic, biochemical, and serological characterization of a new pneumococcal serotype, 6H, and generation of a pneumococcal strain producing three different capsular repeat units. *Clin Vaccine Immunol* 22:313–318. <https://doi.org/10.1128/CVI.00647-14>.
 57. Kjeldsen C, Slott S, Elverdal PL, Sheppard CL, Kapatai G, Fry NK, Skovsted IC, Duus JØ. 2018. Discovery and description of a new serogroup 7 *Streptococcus pneumoniae* serotype, 7D, and structural analysis of 7C and 7D. *Carbohydr Res* 463:24–31. <https://doi.org/10.1016/j.carres.2018.04.011>.
 58. Sorensen UB. 1993. Typing of pneumococci by using 12 pooled antisera. *J Clin Microbiol* 31:2097–2100. <https://doi.org/10.1128/JCM.31.8.2097-2100.1993>.